DOCUMENT RESUME

ED 045 732 TM 000 322

| | |
|---|---|
| AUTHOR | Wender, Karl |
| TITLE | Independence of Dimensions in Multidimensional Scaling. |
| INSTITUTION | Technische Hochschule Darmstadt (West Germany). |
| PUB DATE | Jun 70 |
| NOTE | 19p. |
| EDRS PRICE | EDRS Price MF-$0.25 HC-$1.05 |
| DESCRIPTORS | Geometric Concepts, *Mathematical Models, *Research Methodology, *Statistical Analysis, *Topology, Visual Discrimination, *Visual Perception |
| IDENTIFIERS | *Multidimensional Scaling |

ABSTRACT

Models for multidimensional scaling use metric spaces with additive difference metrics. Two important properties of additive difference metrics are decomposability and intradimensional subtractivity. A prediction was derived from these properties and tested experimentally. Eleven non-psychology students were used as subjects. Rectangles varying in area and shape served as stimuli. Two types of dissimilarity judgments were obtained: rating scale judgment and pair comparison. In the rating scale experiment two stimuli at a time were shown and the students were asked to judge the overall difference on a scale running from 0 for "no difference" to 8 for "very great difference." In the pair comparisons, slides showed two pairs and the subjects had to judge which of the two pairs showed the larger overall difference. The decompcsability and intradimensional subtractivity assumptions of the model were violated by most students due to an apparent interaction between area and shape. The experiment failed to confirm the model and it is suggested that the model be rejected or modified. (Author/LR)

# Independence of Dimensions in Multidimensional Scaling

Karl Wender

D a r m s t a d t

Juni 1970

INDEPENDENCE OF DIMENSIONS IN MULTIDIMENSIONAL SCALING

Karl Wender

Technische Hochschule Darmstadt

Abstract

Models for multidimensional scaling use metric spaces with additive difference metrics. Two important properties of additive difference metrics are decomposability and intradimensional subtractivity. A prediction is derived from these properties and tested experimentally. Rectangles varying in area and shape served as stimuli. Dissimilarity judgments were obtained by both rating and pair comparison procedures. The assumptions of the model are violated by most of the Ss. Apparently this violation is due to an interaction between the two dimensions.

# INDEPENDENCE OF DIMENSIONS IN MULTIDIMENSIONAL SCALING

Karl Wender

Technische Hochschule Darmstadt

In recent years models for multidimensional scaling have been advanced in several ways. One of the major contributions is that of Beals and Krantz (1967), Beals, Krantz, and Tversky (1968), and Tversky and Krantz (1969), who introduced a system of axioms for ordinal multidimensional scaling and proved appropriate representation and uniqueness theorems. In particular these axioms show how restrictive the assumptions are on which commonly used multidimensional scaling procedures are based. Not all of the axioms are testable in the sense of Pfanzagl (1968, p. 108). But at least some of them allow one to design an experiment which eventually may show that the axiom does not hold under particular circumstances.

So far there are few empirical investigations of the axioms. The only published study known to the present author is that of Tversky and Krantz (1969) who tested predictions derived from the axioms for interdimensional additivity. Using schematic faces as stimuli Tversky and Krantz inferred from their data that interdimensional additivity might well be satisfied but as they noted this conclusion is valid only for their type of stimuli and configuration. In fact the study involved only very few stimuli presented in a special, regular configuration. This configuration may have biased their results toward not violating the assumptions.

The present study investigates a similar property of the model using different stimuli and a different configuration. Furthermore, the hypothesis being tested is that with the same kind of stimuli but different configurations the model might be violated in different ways due to context effects.

Metric spaces commonly in use with multidimensional scaling procedures are characterized by additive segments and additive difference metrics. This paper is concerned with a special property of additive difference metrics.

Beals, Krantz, and Tversky (1968) write the general model of additive difference metrics in the following way

$$(1)\qquad d(x,y) = F\left(\sum_{i}^{n} \Phi_{i}\left(|X_i - Y_i|\right)\right)$$

$d(x,y)$ is the subjective difference between two stimuli $x$ and $y$. The $x_i$ and $y_i$ are their values on the "relevant" physical dimensions. $X_i = f_i(x_i)$ and $Y_i = f_i(y_i)$ are the coordinates of the corresponding points in subjective space. The $f_i$ are real valued functions, sometimes called psychophysical functions. The essence of the models lies in three properties: (1) decomposability; (2) interdimensional additivity; and (3); intradimensional subtractivity. <u>Decomposability</u> requires that there be no interaction between the dimensions of the subjective space i.e. the dimensions contribute independently to the overall distance. In equation (1) this is expressed by the fact that the arguments of the $\Phi$ functions for each i belong to the same i'th dimension. Intradimensional subtractivity specifies that on each dimension the absolute value of the difference between the corresponding coordinates is computed. To satisfy interdimensional additivity these contributions are combined by addition after some monotone transformation $\Phi_i$. A further monotone transformation F gives the distance between the points.

The present paper is primarily concerned with decomposability and subtractivity. The general model for these properties may be written as

$$(2) \qquad d(x,y) = F(\Phi_1(|X_1-Y_1|),\ldots, \Phi_n(|X_n-Y_n|))$$

where all symbols are defined as before except F which is now a function of n variables.

From equation (2) follows immediately: If x, y, u, v are stimuli with $x_i = y_i$; $u_i = v_i$ for some i and with $x_j = u_j$; $y_j = v_j$ for $j = 1,\ldots, n$, $j \neq i$; then: $d(x,y) = d(u,v)$. This property of the model is the one being tested in the following experiment.

In an unpublished experiment the present author found that a multidimensional scaling study using the complete method of triads (Torgerson, 1958) with rectangles varying in shape and area resulted in an apparent interaction between these two dimensions. That is, subjective dissimilarities which should be due only to differences in area seemed to depend also on differences in shape. This is a violation of decomposability. Furthermore, the data suggested that dissimilarity judgments in some region of the space might be influenced by the degree to which this region is represented in the set of stimuli used throughout the experiment. If one region in the space is represented by relative many stimuli, Ss may respond to them by making larger dissimilarity judgments as compared with judgments about stimuli from less well represented regions. Somewhat similar results have been found in unidimensional scaling, e.g. Stevens, 1959. Since the study mentioned above was conducted for a different purpose the present experiment is designed to investigate this hypothesis more accurately.

## Method

The stimuli used were rectangles varying in area and shape. Shape was defined as the ratio of height to width. As shown in figure 1 three sets of stimuli were prepared.

- - - - - - - - - - - - - - - -

Insert figure 1 about here

- - - - - - - - - - - - - - - -

Stimuli 1 through 10 were common to all three sets (a), (b), and (c). Sets (b) and (c) contained additional stimuli, which were more like squares in set (b) and more elongated in set (c).

The rectangles were cut from heavy white paper, photographed, and shown by a slide projector. They appeared as dark figures on a bright background.

Two types of dissimilarity judgments were obtained: rating scale judgments and pair comparisons. In the rating scale experiment two stimuli at a time were shown and Ss were asked to judge the overall difference on a rating scale running from 0 for "no difference" to 8 for "very large difference". The following pairs of stimuli were prepared for display.

set (a): (1,6), (2,7), (3,8), (4,9), (5,10)
set (b): same as for (a) plus (11,12), (15,16),
(19,20), (23,24),
set (c): same as for (a) plus (13,14), (17,18),
(21,22), (25,26).

The two rectangles of each pair appeared on the slides in upright position with their midpoints at the same height and separated by a distance which was about 3 times their width. Each stimulus pair was photographed 6 times; 3 times each in reversed left-right position. This gave 30 slides for experiment (a) and 54 slides for experiments (b) and (c). The slides were presented in random sequence.

Each of the experiments began with 30 training trials for which data were excluded from the analysis. During the experiments proper each slide was judged twice by all Ss resulting in 12 judgments on each stimulus pair by each S.

For the pair comparison experiment the same stimulus pairs were used as above but now all possible pairs of pairs were presented. The slides showed 4 stimuli: one pair on the left side and one on the right side separated by a dark line. The Ss had to judge which of the two pairs showed the larger overall difference.

For experiment (d) all 10 possible combinations of pairs from set (a) were photographed 6 times, left-right position counterbalanced. The 60 slides were presented in random sequence and judged twice so that again 12 judgments on each stimulus combination were obtained. Sets (e) and (f) were made from sets (b) and (c) respectively. From the 9 pairs 36 quadruples were photographed twice, left-right position counterbalanced. This series of 72 slides was judged 6 times by each S, resulting in 12 judgments on all stimulus combinations.

The Ss were 11 non-psychology students who were paid for their participation. Each S was given all sets (a) through (e) in individual sessions. The sessions for set (a) through (d) took 30 to 45 minutes. Experiments (e) and (f) were divided into three parts, each of which took the same time. Each S had a total of 10 sessions on 10 different days, each session starting with 30 training trials. Slide presentation was controlled by an electronic timing device. Slides were shown by a slide projector for 8 seconds during which time the Ss responded by pressing a button. The judgments were punched on a paper tape.

## Results

The data were analysed separately for each S. Only those stimulus pairs which were common to all six subexperiments were included in the analysis, i.e. the pairs (1,6) through (5,10).

Individual data for the rating scale experiment consisted of a 5 by 12 matrix with integers between 0 and 8 as entries. If equation (2) holds the rowmeans should vary only due to random fluctuations. To test equation (2) the Kruskal Wallis H statistic was computed since ordinal multidimensional scaling requires only ordinal scale data. In table 1 the resulting $\chi^2$ values for each S are given together with the arithmetic means of the difference judgments for each of the stimulus pairs. All $\chi^2$ values exceeding 15,3 (indicated by an asterisk) are significant at the .01 level.

- - - - - - - - - - - - - -

Insert table 1 about here

- - - - - - - - - - - - - -

In the pair comparison experiment each slide showed one pair of stimuli that looked more like squares, the other pair being relatively slender. If a S chose the pair consisting of the thinner rectangles as having the larger overall difference this judgment was scored "1", and "0" otherwise. Thus the data for each S consist of a 5 by 5 pair comparison matrix with the main diagonal left empty. The entries of the matrix show how often the thinner pair was chosen over the more square one. The sum of the elements in this matrix was used to test the model. The sum equals 120 if a S chooses the thinner pair on every trial and it becomes 0 if a S always chooses the more square rectangles. Under the hypothesis that S has a preference probability of .5 for the more square stimuli the sum should follow a binomial distribution with a mean of 60.

Under this assumption the sum of the matrix elements lies between 45 and 75 with probability .99. If it reaches or exceeds these boundaries the hypothesis is rejected at the .01 level of significance.

To show the trend of a possible violation it was determined for each pair how often it was chosen over any other pair. These values together with the sums of the pair comparison matrices are given in table 2.

- - - - - - - - - - - - - -

Insert table 2 about here

- - - - - - - - - - - - - -

## Discussion

On the basis of the rating scale data the multidimensional scaling model is to be rejected for about half the Ss. Almost all Ss violated the model in the pair comparison experiment. The difference between the rating scale and the pair comparison data may be due to some methodological artifact. For example, the rating scale data may contain a larger amount of error.

As is shown in tables 1 and 2 most of the Ss judge the overal difference to be larger for thin rectangles and to be smaller for squares. Furthermore, there is no systematic difference between the subexperiments. Thus the hypothesis regarding context effects is not supported. For most Ss the data appear very consistent. For some Ss all 120 judgments of the pair comparison experiment are in the same direction. In a very few cases judgments vary in an irregular manner, e.g. difference judgments being small for squares and thin rectangles and being larger for rectangles in between.

The main result of this study is that the prediction derived from decomposability and intradimensional subtractivity are violated by most of the Ss. Of course it is possible that the violation of the model was caused by the special selection of stimuli and by the fact that not all possible pairs were presented throughout the experiment. Most probably Ss were aware of this and, accordingly, some response bias may have been introduced. Unfortunately, there is no way to control response bias in these types of scaling methods, and the multidimensional scaling model might hold for the same type of stimuli in different experimental conditions.

Though there remains the puzzling result that the multidimensional scaling model was violated by stimuli as simple as rectangles varying in area and shape we cannot conclude that this will be the case when the model is applied to more complex stimuli. The stimuli used here are of the kind called analyzable by Shepard (1964). As suggested by Torgerson (1965) multidimensional scaling models might be more appropriate for unitary stimuli.

It is possible, at least in principle, that the model may be satisfied with the same stimuli but with physical variables other than area and shape. Two obvious alternatives are height and width of rectangles. The unpublished study mentioned above, however, suggested that area and shape are the more relevant dimensions.

After this failure to confirm the model of multidimensional scaling two alternatives remain. One is to completely reject the model and the other is to modify it. One modification that comes to mind is to redefine the psychophysical functions to include interactions between the dimensions. But from the point of view of economy such a complication of the model seems to be undesirable.

## References


Beals, R., & Krantz, D.H. Metrics and geodesics induced by order relations. Mathematische Zeitschrift, 1967, 101, 285-298.

Beals, R., Krantz, D.H., & Tversky, A. Foundations of multidimensional scaling. Psychological Review, 1968, 75, 127-142.

Pfanzagl, J. Theory of measurement. Würzburg: Physica, 1968.

Shepard, R.N. Attention and the metric structure of stimulus space. Journal of Mathematical Psychology, 1964, 1, 54-87.

Stevens, S.S. Measurement, psychophysics and utility. In: Churchman, C.W., & P. Ratoosh (Eds.), Measurement definitions and theories. New York: Wiley, 1959, Pp. 18-63.

Torgerson, W.S. Theory and method of scaling. New York: Wiley, 1958.

Torgerson, W.S. Multidimensional scaling of similarity. Psychometrika, 1965, 30, 379-393.

Tversky, A., & Krantz, D.H. Similarity of schematic faces: A test of interdimensional additivity. Perception & Psychophysics, 1969, 5, 124-128.

Table 1

Mean Ratings of Difference and $\chi^2$ Values

| Subject | Stimulus Pair | | | | | $\chi^2$ |
|---|---|---|---|---|---|---|
| | 1,6 | 2,7 | 3,8 | 4,9 | 5,1o | |
| | | | Set (a) | | | |
| 1 | 3.4 | 3.6 | 5.3 | 6.o | 7.o | 23.9 * |
| 2 | 5.3 | 5.1 | 4.7 | 5.9 | 6.6 | 7.5 |
| 3 | 4.3 | 5.5 | 5.o | 5.3 | 6.2 | 3.8 |
| 4 | 4.o | 4.2 | 5.7 | 6.4 | 6.2 | 31.o * |
| 5 | 5.1 | 4.8 | 5.3 | 4.8 | 4.4 | 3.7 |
| 6 | | missing data | | | | |
| 7 | 2.3 | 3.1 | 5.9 | 6.6 | 6.3 | 28.8 * |
| 8 | 3.3 | 4.1 | 4.5 | 5.3 | 5.5 | 29.2 * |
| 9 | 1.5 | 3.7 | 4.4 | 7.3 | 7.7 | 44.5 * |
| 1o | 3.7 | 4.2 | 5.8 | 5.1 | 5.3 | 8.1 |
| 11 | 4.o | 4.3 | 4.7 | 4.6 | 5.5 | 6.o |
| | | | Set (b) | | | |
| 1 | 1.6 | 2.7 | 2.5 | 2.8 | 2.6 | 12.8 |
| 2 | 1.o | 2.1 | 2.3 | 2.4 | 2.7 | 3o.5 * |
| 3 | 2.6 | 3.3 | 3.1 | 2.8 | 3.3 | 2.7 |
| 4 | 2.8 | 2.8 | 3.1 | 2.6 | 3.o | 3.8 |
| 5 | 1.7 | 2.3 | 2.9 | 3.1 | 3.7 | 23.4 * |
| 6 | 1.2 | 2.1 | 2.o | 2.8 | 2.8 | 27.2 * |
| 7 | 1.5 | 2.o | 2.2 | 1.8 | 2.2 | 1.6 |
| 8 | 3.3 | 3.9 | 3.9 | 3.o | 3.2 | 15.6 * |
| 9 | 1.8 | 2.3 | 1.8 | 1.7 | 2.3 | 2.7 |
| 1o | 1.8 | 1.8 | 1.9 | 2.1 | 2.8 | 8.2 |
| 11 | 1.9 | 2.3 | 2.5 | 2.5 | 2.5 | 5.o |

Table 1 (cont.)

Set (c)

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 2.2 | 2.2 | 2.3 | 2.3 | 2.3 | 0.8 |
| 2 | 1.2 | 1.8 | 1.8 | 1.8 | 1.9 | 5.4 |
| 3 | 2.3 | 3.4 | 3.2 | 3.4 | 3.4 | 10.2 |
| 4 | 2.6 | 2.6 | 2.8 | 2.7 | 2.8 | 1.0 |
| 5 | 1.4 | 1.8 | 2.1 | 2.3 | 2.0 | 6.7 |
| 6 | 1.7 | 2.0 | 2.1 | 2.1 | 2.3 | 4.0 |
| 7 | 1.4 | 2.4 | 2.5 | 2.4 | 2.3 | 9.7 |
| 8 | 3.7 | 4.8 | 5.0 | 4.8 | 3.7 | 23.1 * |
| 9 | 1.4 | 3.3 | 3.3 | 2.7 | 2.3 | 10.8 |
| 10 | 2.3 | 1.9 | 1.8 | 1.9 | 1.8 | 2.0 |
| 11 | 2.5 | 2.5 | 2.2 | 2.3 | 2.3 | 2.3 |

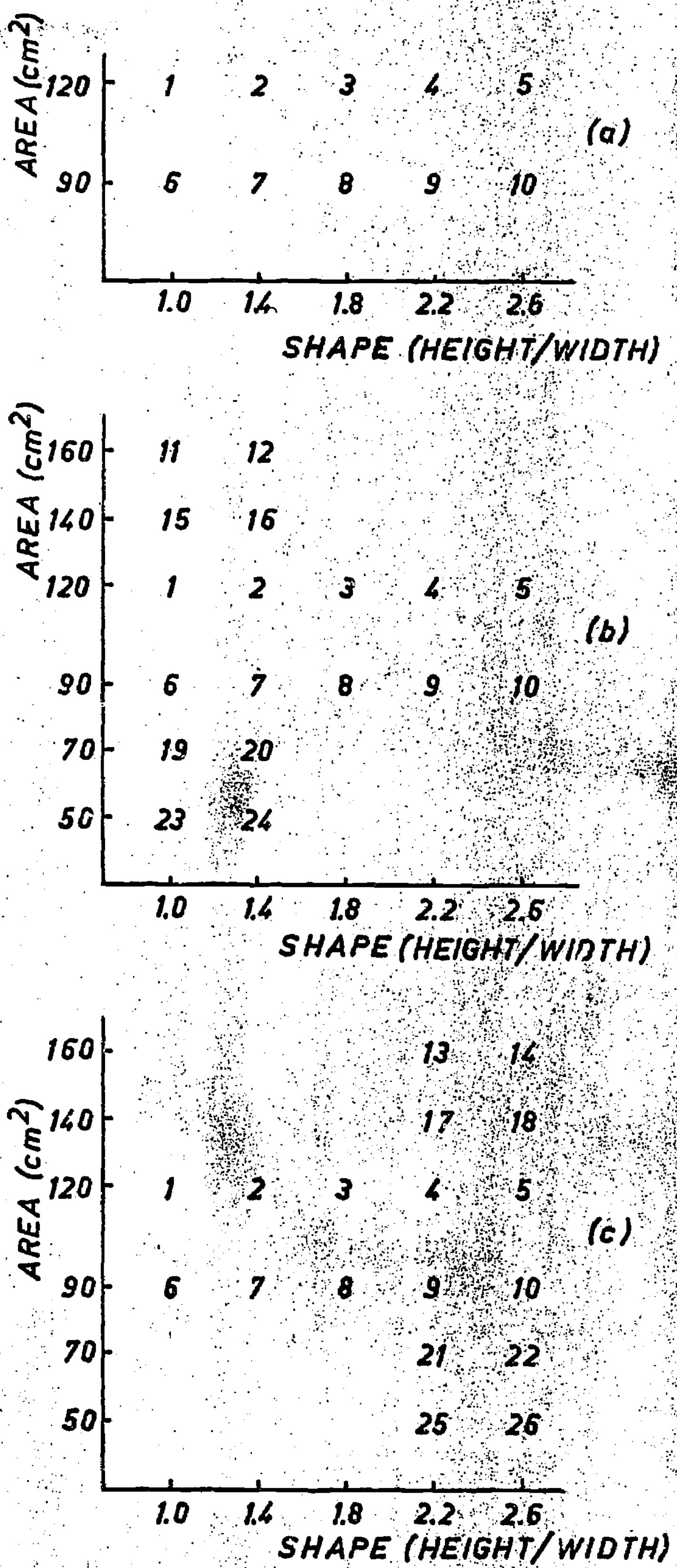
Table 2

Number of Times Each Stimulus Pair Was Preferred to Any Other Pair and Sums of the Pair Comparison Matrices

| Subject | Stimulus Pair | | | | | |
|---|---|---|---|---|---|---|
| | 1,6 | 2,7 | 3,8 | 4,9 | 5,1o | |
| | | | Set (d) | | | |
| 1 | o | 14 | 24 | 34 | 48 | 118 * |
| 2 | 34 | 28 | 2o | 19 | 19 | 43 * |
| 3 | 45 | 26 | 22 | 17 | 1o | 25 * |
| 4 | 1 | 12 | 25 | 34 | 48 | 118 * |
| 5 | 3 | 11 | 25 | 34 | 47 | 116 * |
| 6 | 47 | 24 | 19 | 17 | 13 | 29 * |
| 7 | 34 | 29 | 24 | 18 | 15 | 41 * |
| 8 | o | 12 | 24 | 36 | 48 | 12o * |
| 9 | 7 | 18 | 25 | 31 | 39 | 97 * |
| 1o | 44 | 33 | 22 | 9 | 12 | 23 * |
| 11 | 28 | 21 | 2o | 26 | 25 | 59 |
| | | | Set (e) | | | |
| 1 | o | 12 | 24 | 36 | 48 | 12o * |
| 2 | 4 | 18 | 28 | 32 | 38 | 94 * |
| 3 | 18 | 32 | 32 | 22 | 16 | 52 |
| 4 | o | 17 | 26 | 36 | 41 | 11o * |
| 5 | 24 | 34 | 33 | 14 | 15 | 45 * |
| 6 | 21 | 18 | 22 | 33 | 26 | 73 |
| 7 | 1o | 19 | 27 | 29 | 35 | 87 * |
| 8 | 2 | 13 | 22 | 35 | 48 | 117 * |
| 9 | 9 | 17 | 22 | 36 | 36 | 99 * |
| 1o | 37 | 22 | 21 | 16 | 24 | 47 |
| 11 | 2 | 19 | 25 | 31 | 43 | 1o3 * |
| | | | Set (f) | | | |
| 1 | o | 13 | 25 | 34 | 47 | 118 * |
| 2 | 3o | 28 | 27 | 19 | 16 | 5o |
| 3 | 33 | 18 | 26 | 25 | 18 | 56 |
| 4 | o | 12 | 24 | 36 | 48 | 12o * |
| 5 | 17 | 13 | 26 | 29 | 35 | 87 * |
| 6 | 45 | 29 | 19 | 16 | 11 | 27 * |
| 7 | 16 | 33 | 3o | 25 | 16 | 55 |
| 8 | o | 12 | 24 | 37 | 47 | 119 * |
| 9 | 5 | 14 | 27 | 36 | 38 | 1o2 * |
| 1o | 47 | 31 | 21 | 15 | 6 | 14 * |
| 11 | 2o | 14 | 25 | 29 | 32 | 76 * |

## Figure Caption

Figure 1. Stimulus configurations used in sets (a), (b), and (c).



AREA (cm²)
120
90
1 2 3 4 5
6 7 8 9 10
(a)
1.0 1.4 1.8 2.2 2.6
SHAPE (HEIGHT/WIDTH)
AREA (cm²)
160
140
120
90
70
50
11 12
15 16
1 2 3 4 5
6 7 8 9 10
19 20
23 24
(b)
1.0 1.4 1.8 2.2 2.6
SHAPE (HEIGHT/WIDTH)
AREA (cm²)
160
140
120
90
70
50
13 14
17 18
1 2 3 4 5
6 7 8 9 10
21 22
25 26
(c)
1.0 1.4 1.8 2.2 2.6
SHAPE (HEIGHT/WIDTH)

# Independence of Dimensions in Multidimensional Scaling

Karl Wender

D a r m s t a d t

Juni 1970

# INDEPENDENCE OF DIMENSIONS IN MULTIDIMENSIONAL SCALING

Karl Wender

Technische Hochschule Darmstadt

## Abstract

Models for multidimensional scaling use metric spaces with additive difference metrics. Two important properties of additive difference metrics are decomposability and intradimensional subtractivity. A prediction is derived from these properties and tested experimentally. Rectangles varying in area and shape served as stimuli. Dissimilarity judgments were obtained by both rating and pair comparison procedures. The assumptions of the model are violated by most of the Ss. Apparently this violation is due to an interaction between the two dimensions.

# INDEPENDENCE OF DIMENSIONS IN MULTIDIMENSIONAL SCALING

Karl Wender

Technische Hochschule Darmstadt

In recent years models for multidimensional scaling have been advanced in several ways. One of the major contributions is that of Beals and Krantz (1967), Beals, Krantz, and Tversky (1968), and Tversky and Krantz (1969), who introduced a system of axioms for ordinal multidimensional scaling and proved appropriate representation and uniqueness theorems. In particular these axioms show how restrictive the assumptions are on which commonly used multidimensional scaling procedures are based. Not all of the axioms are testable in the sense of Pfanzagl (1968, p. 108). But at least some of them allow one to design an experiment which eventually may show that the axiom does not hold under particular circumstances.

So far there are few empirical investigations of the axioms. The only published study known to the present author is that of Tversky and Krantz (1969) who tested predictions derived from the axioms for interdimensional additivity. Using schematic faces as stimuli Tversky and Krantz inferred from their data that interdimensional additivity might well be satisfied but as they noted this conclusion is valid only for their type of stimuli and configuration. In fact the study involved only very few stimuli presented in a special, regular configuration. This configuration may have biased their results toward not violating the assumptions.

The present study investigates a similar property of the model using different stimuli and a different configuration. Furthermore, the hypothesis being tested is that with the same kind of stimuli but different configurations the model might be violated in different ways due to context effects.

Metric spaces commonly in use with multidimensional scaling procedures are characterized by additive segments and additive difference metrics. This paper is concerned with a special property of additive difference metrics.

Beals, Krantz, and Tversky (1968) write the general model of additive difference metrics in the following way

$$(1) \qquad d(x,y) = F(\sum_{i}^{n} \Phi_{i} (|X_i - Y_i|))$$

$d(x,y)$ is the subjective difference between two stimuli $x$ and $y$. The $x_i$ and $y_i$ are their values on the "relevant" physical dimensions. $X_i = f_i(x_i)$ and $Y_i = f_i(y_i)$ are the coordinates of the corresponding points in subjective space. The $f_i$ are real valued functions, sometimes called psychophysical functions. The essence of the models lies in three properties: (1) decomposability; (2) interdimensional additivity; and (3); intradimensional subtractivity. <u>Decomposability</u> requires that there be no interaction between the dimensions of the subjective space i.e. the dimensions contribute independently to the overall distance. In equation (1) this is expressed by the fact that the arguments of the $\Phi$ functions for each i belong to the same i'th dimension. Intradimensional subtractivity specifies that on each dimension the absolute value of the difference between the corresponding coordinates is computed. To satisfy interdimensional additivity these contributions are combined by addition after some monotone transformation $\Phi_i$. A further monotone transformation F gives the distance between the points.

The present paper is primarily concerned with decomposability and subtractivity. The general model for these properties may be written as

$$(2) \qquad d(x,y) = F(\Phi_1(|X_1-Y_1|),\ldots, \Phi_n(|X_n-Y_n|))$$

where all symbols are defined as before except F which is now a function of n variables.

From equation (2) follows immediately: If x, y, u, v are stimuli with $x_i = y_i$; $u_i = v_i$ for some i and with $x_j = u_j$; $y_j = v_j$ for $j = 1,\ldots, n$, $j \neq i$; then: $d(x,y) = d(u,v)$. This property of the model is the one being tested in the following experiment.

In an unpublished experiment the present author found that a multidimensional scaling study using the complete method of triads (Torgerson, 1958) with rectangles varying in shape and area resulted in an apparent interaction between these two dimensions. That is, subjective dissimilarities which should be due only to differences in area seemed to depend also on differences in shape. This is a violation of decomposability. Furthermore, the data suggested that dissimilarity judgments in some region of the space might be influenced by the degree to which this region is represented in the set of stimuli used throughout the experiment. If one region in the space is represented by relative many stimuli, Ss may respond to them by making larger dissimilarity judgments as compared with judgments about stimuli from less well represented regions. Somewhat similar results have been found in unidimensional scaling, e.g. Stevens, 1959. Since the study mentioned above was conducted for a different purpose the present experiment is designed to investigate this hypothesis more accurately.

## Method

The stimuli used were rectangles varying in area and shape. Shape was defined as the ratio of height to width. As shown in figure 1 three sets of stimuli were prepared.

- - - - - - - - - - - - - - - - -

Insert figure 1 about here

- - - - - - - - - - - - - - - - -

Stimuli 1 through 10 were common to all three sets (a), (b), and (c). Sets (b) and (c) contained additional stimuli, which were more like squares in set (b) and more elongated in set (c).

The rectangles were cut from heavy white paper, photographed, and shown by a slide projector. They appeared as dark figures on a bright background.

Two types of dissimilarity judgments were obtained: rating scale judgments and pair comparisons. In the rating scale experiment two stimuli at a time were shown and Ss were asked to judge the overall difference on a rating scale running from 0 for "no difference" to 8 for "very large difference". The following pairs of stimuli were prepared for display.

set (a): (1,6), (2,7), (3,8), (4,9), (5,10)
set (b): same as for (a) plus (11,12), (15,16),
(19,20), (23,24),
set (c): same as for (a) plus (13,14), (17,18),
(21,22), (25,26).

The two rectangles of each pair appeared on the slides in upright position with their midpoints at the same height and separated by a distance which was about 3 times their width. Each stimulus pair was photographed 6 times; 3 times each in reversed left-right position. This gave 30 slides for experiment (a) and 54 slides for experiments (b) and (c). The slides were presented in random sequence.

Each of the experiments began with 30 training trials for which data were excluded from the analysis. During the experiments proper each slide was judged twice by all Ss resulting in 12 judgments on each stimulus pair by each S.

For the pair comparison experiment the same stimulus pairs were used as above but now all possible pairs of pairs were presented. The slides showed 4 stimuli: one pair on the left side and one on the right side separated by a dark line. The Ss had to judge which of the two pairs showed the larger overall difference.

For experiment (d) all 10 possible combinations of pairs from set (a) were photographed 6 times, left-right position counterbalanced. The 60 slides were presented in random sequence and judged twice so that again 12 judgments on each stimulus combination were obtained. Sets (e) and (f) were made from sets (b) and (c) respectively. From the 9 pairs 36 quadruples were photographed twice, left-right position counterbalanced. This series of 72 slides was judged 6 times by each S, resulting in 12 judgments on all stimulus combinations.

The Ss were 11 non-psychology students who were paid for their participation. Each S was given all sets (a) through (e) in individual sessions. The sessions for set (a) through (d) took 30 to 45 minutes. Experiments (e) and (f) were divided into three parts, each of which took the same time. Each S had a total of 10 sessions on 10 different days, each session starting with 30 training trials. Slide presentation was controlled by an electronic timing device. Slides were shown by a slide projector for 8 seconds during which time the Ss responded by pressing a button. The judgments were punched on a paper tape.

## Results

The data were analysed separately for each S. Only those stimulus pairs which were common to all six sub-experiments were included in the analysis, i.e. the pairs (1,6) through (5,10).

Individual data for the rating scale experiment consisted of a 5 by 12 matrix with integers between 0 and 8 as entries. If equation (2) holds the rowmeans should vary only due to random fluctuations. To test equation (2) the Kruskal Wallis H statistic was computed since ordinal multidimensional scaling requires only ordinal scale data. In table 1 the resulting $\chi^2$ values for each S are given together with the arithmetic means of the difference judgments for each of the stimulus pairs: All $\chi^2$ values exceeding 15,3 (indicated by an asterisk) are significant at the .01 level.

- - - - - - - - - - - - - - -

Insert table 1 about here

- - - - - - - - - - - - - - -

In the pair comparison experiment each slide showed one pair of stimuli that looked more like squares, the other pair being relatively slender. If a S chose the pair consisting of the thinner rectangles as having the larger overall difference this judgment was scored "1", and "0" otherwise. Thus the data for each S consist of a 5 by 5 pair comparison matrix with the main diagonal left empty. The entries of the matrix show how often the thinner pair was chosen over the more square one. The sum of the elements in this matrix was used to test the model. The sum equals 120 if a S chooses the thinner pair on every trial and it becomes 0 if a S always chooses the more square rectangles. Under the hypothesis that S has a preference probability of .5 for the more square stimuli the sum should follow a binomial distribution with a mean of 60.

Under this assumption the sum of the matrix elements lies between 45 and 75 with probability .99. If it reaches or exceeds these boundaries the hypothesis is rejected at the .01 level of significance.

To show the trend of a possible violation it was determined for each pair how often it was chosen over any other pair. These values together with the sums of the pair comparison matrices are given in table 2.

- - - - - - - - - - - - -

Insert table 2 about here

- - - - - - - - - - - - -

## Discussion

On the basis of the rating scale data the multidimensional scaling model is to be rejected for about half the Ss. Almost all Ss violated the model in the pair comparison experiment. The difference between the rating scale and the pair comparison data may be due to some methodological artifact. For example, the rating scale data may contain a larger amount of error.

As is shown in tables 1 and 2 most of the Ss judge the overal difference to be larger for thin rectangles and to be smaller for squares. Furthermore, there is no systematic difference between the subexperiments. Thus the hypothesis regarding context effects is not supported. For most Ss the data appear very consistent. For some Ss all 120 judgments of the pair comparison experiment are in the same direction. In a very few cases judgments vary in an irregular manner, e.g. difference judgments being small for squares and thin rectangles and being larger for rectangles in between.

The main result of this study is that the prediction derived from decomposability and intradimensional subtractivity are violated by most of the Ss. Of course it is possible that the violation of the model was caused by the special selection of stimuli and by the fact that not all possible pairs were presented throughout the experiment. Most probably Ss were aware of this and, accordingly, some response bias may have been introduced. Unfortunately, there is no way to control response bias in these types of scaling methods, and the multidimensional scaling model might hold for the same type of stimuli in different experimental conditions.

Though there remains the puzzling result that the multidimensional scaling model was violated by stimuli as simple as rectangles varying in area and shape we cannot conclude that this will be the case when the model is applied to more complex stimuli. The stimuli used here are of the kind called analyzable by Shepard (1964). As suggested by Torgerson (1965) multidimensional scaling models might be more appropriate for unitary stimuli.

It is possible, at least in principle, that the model may be satisfied with the same stimuli but with physical variables other than area and shape. Two obvious alternatives are height and width of rectangles. The unpublished study mentioned above, however, suggested that area and shape are the more relevant dimensions.

After this failure to confirm the model of multidimensional scaling two alternatives remain. One is to completely reject the model and the other is to modify it. One modification that comes to mind is to redefine the psychophysical functions to include interactions between the dimensions. But from the point of view of economy such a complication of the model seems to be undesirable.

## References

Beals, R., & Krantz, D.H. Metrics and geodesics induced by order relations. Mathematische Zeitschrift, 1967, 101, 285-298.

Beals, R., Krantz, D.H., & Tversky, A. Foundations of multidimensional scaling. Psychological Review, 1968, 75, 127-142.

Pfanzagl, J. Theory of measurement. Würzburg: Physica, 1968.

Shepard, R.N. Attention and the metric structure of stimulus space. Journal of Mathematical Psychology, 1964, 1, 54-87.

Stevens, S.S. Measurement, psychophysics and utility. In: Churchman, C.W., & P. Ratoosh (Eds.), Measurement definitions and theories. New York: Wiley, 1959, Pp. 18-63.

Torgerson, W.S. Theory and method of scaling. New York: Wiley, 1958.

Torgerson, W.S. Multidimensional scaling of similarity. Psychometrika, 1965, 30, 379-393.

Tversky, A., & Krantz, D.H. Similarity of schematic faces: A test of interdimensional additivity. Perception & Psychophysics, 1969, 5, 124-128.

Table 1

Mean Ratings of Difference and $\chi^2$ Values

| Subject | 1,6 | 2,7 | 3,8 | 4,9 | 5,1o | $\chi^2$ |
|---|---|---|---|---|---|---|
| | | | Set (a) | | | |
| 1 | 3.4 | 3.6 | 5.3 | 6.o | 7.o | 23.9 * |
| 2 | 5.3 | 5.1 | 4.7 | 5.9 | 6.6 | 7.5 |
| 3 | 4.3 | 5.5 | 5.o | 5.3 | 6.2 | 3.8 |
| 4 | 4.o | 4.2 | 5.7 | 6.4 | 6.2 | 31.o * |
| 5 | 5.1 | 4.8 | 5.3 | 4.8 | 4.4 | 3.7 |
| 6 | | missing data | | | | |
| 7 | 2.3 | 3.1 | 5.9 | 6.6 | 6.3 | 28.8 * |
| 8 | 3.3 | 4.1 | 4.5 | 5.3 | 5.5 | 29.2 * |
| 9 | 1.5 | 3.7 | 4.4 | 7.3 | 7.7 | 44.5 * |
| 1o | 3.7 | 4.2 | 5.8 | 5.1 | 5.3 | 8.1 |
| 11 | 4.o | 4.3 | 4.7 | 4.6 | 5.5 | 6.o |
| | | | Set (b) | | | |
| 1 | 1.6 | 2.7 | 2.5 | 2.8 | 2.6 | 12.8 |
| 2 | 1.o | 2.1 | 2.3 | 2.4 | 2.7 | 3o.5 * |
| 3 | 2.6 | 3.3 | 3.1 | 2.8 | 3.3 | 2.7 |
| 4 | 2.8 | 2.8 | 3.1 | 2.6 | 3.o | 3.8 |
| 5 | 1.7 | 2.3 | 2.9 | 3.1 | 3.7 | 23.4 * |
| 6 | 1.2 | 2.1 | 2.o | 2.8 | 2.8 | 27.2 * |
| 7 | 1.5 | 2.o | 2.2 | 1.8 | 2.2 | 1.6 |
| 8 | 3.3 | 3.9 | 3.9 | 3.o | 3.2 | 15.6 * |
| 9 | 1.8 | 2.3 | 1.8 | 1.7 | 2.3 | 2.7 |
| 1o | 1.8 | 1.8 | 1.9 | 2.1 | 2.8 | 8.2 |
| 11 | 1.9 | 2.3 | 2.5 | 2.5 | 2.5 | 5.o |

Table 1 (cont.)

Set (c)

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 2.2 | 2.2 | 2.3 | 2.3 | 2.3 | o.8 |
| 2 | 1.2 | 1.8 | 1.8 | 1.8 | 1.9 | 5.4 |
| 3 | 2.3 | 3.4 | 3.2 | 3.4 | 3.4 | 1o.2 |
| 4 | 2.6 | 2.6 | 2.8 | 2.7 | 2.8 | 1.o |
| 5 | 1.4 | 1.8 | 2.1 | 2.3 | 2.o | 6.7 |
| 6 | 1.7 | 2.o | 2.1 | 2.1 | 2.3 | 4.o |
| 7 | 1.4 | 2.4 | 2.5 | 2.4 | 2.3 | 9.7 |
| 8 | 3.7 | 4.8 | 5.o | 4.8 | 3.7 | 23.1 * |
| 9 | 1.4 | 3.3 | 3.3 | 2.7 | 2.3 | 1o.8 |
| 1o | 2.3 | 1.9 | 1.8 | 1.9 | 1.8 | 2.o |
| 11 | 2.5 | 2.5 | 2.2 | 2.3 | 2.3 | 2.3 |

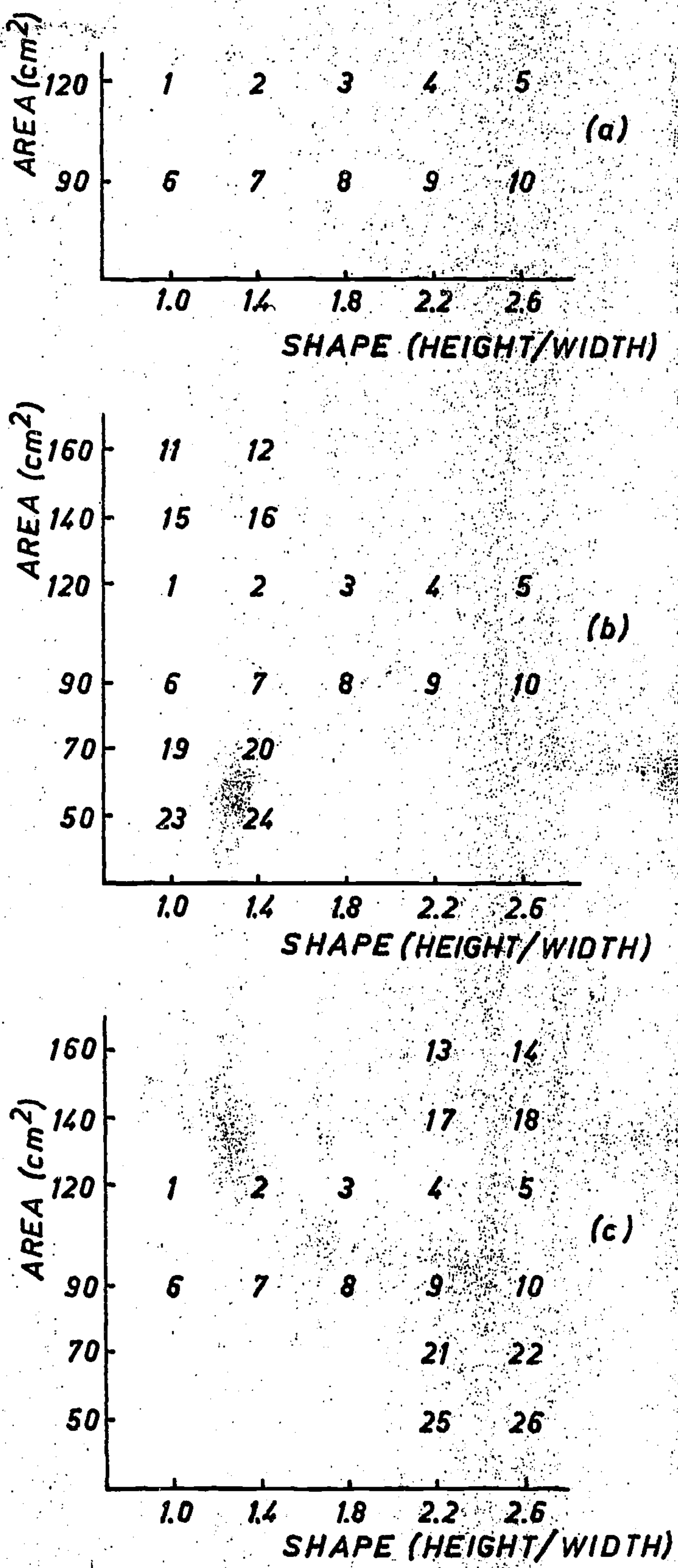
Table 2

Number of Times Each Stimulus Pair Was Preferred to Any Other Pair and Sums of the Pair Comparison Matrices

| Subject | Stimulus Pair 1,6 | 2,7 | 3,8 | 4,9 | 5,1o | | |
|---|---|---|---|---|---|---|---|
| **Set (d)** | | | | | | | |
| 1 | o | 14 | 24 | 34 | 48 | 118 | * |
| 2 | 34 | 28 | 2o | 19 | 19 | 43 | * |
| 3 | 45 | 26 | 22 | 17 | 1o | 25 | * |
| 4 | 1 | 12 | 25 | 34 | 49 | 118 | * |
| 5 | 3 | 11 | 25 | 34 | 47 | 116 | * |
| 6 | 47 | 24 | 19 | 17 | 13 | 29 | * |
| 7 | 34 | 29 | 24 | 18 | 15 | 41 | * |
| 8 | o | 12 | 24 | 36 | 48 | 12o | * |
| 9 | 7 | 18 | 25 | 31 | 39 | 97 | * |
| 1o | 44 | 33 | 22 | 9 | 12 | 23 | * |
| 11 | 28 | 21 | 2o | 26 | 25 | 59 | |
| **Set (e)** | | | | | | | |
| 1 | o | 12 | 24 | 36 | 48 | 12o | * |
| 2 | 4 | 18 | 28 | 32 | 38 | 94 | * |
| 3 | 18 | 32 | 32 | 22 | 16 | 52 | |
| 4 | o | 17 | 26 | 36 | 41 | 11o | * |
| 5 | 24 | 34 | 33 | 14 | 15 | 45 | * |
| 6 | 21 | 18 | 22 | 33 | 26 | 73 | |
| 7 | 1o | 19 | 27 | 29 | 35 | 87 | * |
| 8 | 2 | 13 | 22 | 35 | 48 | 117 | * |
| 9 | 9 | 17 | 22 | 36 | 36 | 99 | * |
| 1o | 37 | 22 | 21 | 16 | 24 | 47 | |
| 11 | 2 | 19 | 25 | 31 | 43 | 1o3 | * |
| **Set (f)** | | | | | | | |
| 1 | o | 13 | 25 | 34 | 47 | 118 | * |
| 2 | 3o | 28 | 27 | 19 | 16 | 5o | |
| 3 | 33 | 18 | 26 | 25 | 18 | 56 | |
| 4 | o | 12 | 24 | 36 | 48 | 12o | * |
| 5 | 17 | 13 | 26 | 29 | 35 | 87 | * |
| 6 | 45 | 29 | 19 | 16 | 11 | 27 | * |
| 7 | 16 | 33 | 3o | 25 | 16 | 55 | |
| 8 | o | 12 | 24 | 37 | 47 | 119 | * |
| 9 | 5 | 14 | 27 | 36 | 38 | 1o2 | * |
| 1o | 47 | 31 | 21 | 15 | 6 | 14 | * |
| 11 | 2o | 14 | 25 | 29 | 32 | 76 | * |

# Figure Caption

Figure 1. Stimulus configurations used in sets (a), (b), and (c).


AREA (cm2)
120
90
1 2 3 4 5
6 7 8 9 10
1.0 1.4 1.8 2.2 2.6
SHAPE (HEIGHT/WIDTH)
(a)
AREA (cm2)
160 140 120 90 70 50
11 12
15 16
1 2 3 4 5
6 7 8 9 10
19 20
23 24
1.0 1.4 1.8 2.2 2.6
SHAPE (HEIGHT/WIDTH)
(b)
AREA (cm2)
160 140 120 90 70 50
13 14
17 18
1 2 3 4 5
6 7 8 9 10
21 22
25 26
1.0 1.4 1.8 2.2 2.6
SHAPE (HEIGHT/WIDTH)
(c)